Research Report CCS-341

MANAGEMENT SCIENCE RELATIONS FOR EVALUATION AND MANAGEMENT ACCOUNTABILITY

by

A. Charnes*
W.W. Cooper**

CENTER FOR CYBERNETIC STUDIES
The University of Texas
Austin, Texas 78712

Research Report CCS-341

MANAGEMENT SCIENCE RELATIONS FOR EVALUATION AND MANAGEMENT ACCOUNTABILITY

by

A. Charnes*
W.W. Cooper**

N00014-75-C-0616
N00014-75-C-0569

May 1979

58

*The University of Texas at Austin
**Harvard University and Carnegie-Mellon University

This paper represents the draft of a paper to be presented as the plenary address on June 20, 1979, for the TIMS XXIV International Meeting in Hawaii. It was partly supported by NSF Grant No. SOC76-15876 Collaborative Research on the Analytical Capabilities of a Goals Accounting System. It is also part of the program of the Navy Personnel Research and Development Center and is partly supported by Projects NR047-071 & 021, ONR Contracts N00014-75-C-0616 & 0569 with the Center for Cybernetic Studies, The University of Texas at Austin, and ONR Contract N00014-76-C-0932 at Carnegie-Mellon University School of Urban and Public Affairs. Reproduction in whole or in part is permitted for any purpose of the U.S. Government.

CENTER FOR CYBERNETIC STUDIES

A. Charnes, Director
Business-Economics Building, 203E
The University of Texas at Austin
Austin, TX 78712
(512) 471-1821

406 197

Abstract

Management Science Relations for Evaluation and Management Accountability

Management accountability as an added dimension for management science research is examined from the standpoint of possible uses in some of the newer "comprehensive auditing" approaches to propriety, effectiveness and efficiency evaluations of management and organization behavior. Attention is centered on non-market activities and not-for-profit organizations. "Goal focussing" is examined, for example, as a relatively recent extension of goal programming for use in effectiveness evaluation and as an alternative to utility theoretic approaches in national goals accounting systems designed to deal with programs or objectives involving numerous kinds of off-market activities. The bulk of the paper, however, is devoted to a new mathematical programming model for deriving analytic representations of extremal frontiers or envelopes from empirical data and for measuring the efficiency of not-for-profit entities. An illustrative application to a recently completed large-scale social experiment in educating disadvantaged children in U.S. public schools is used to show how distinctions may also be drawn between "program efficiency" and "management efficiency." The appendix develops a canonical form for the types of statistical distributions involved. It also provides a beginning for dealing with statistical properties of the extremal relations obtained by applying these kinds of mathematical programming models to observational data.

1. PERSPECTIVES

The occasion of a plenary address can be used in many ways. On this occasion we want to try to provide some perspective on where we have been and where we can (or should) be going. As will soon become clear, we shall mainly reference our own work, but we shall try to do this in a way that relates this work to other major thrusts in the development of management science and/or operations research. Our emphasis will be directed, for the most part, however, toward new paths into the future and improved ways for appraising management activities that might best be summarized in terms such as "accountability" and/or "audit and evaluation."

In broad brush fashion the past history of management science may be idenfified with first an evolution from military to private sector problems. The methodological developments accompanying this evolution ensured that neither the military nor the private sector applications suffered and, indeed, both were strengthended. This continues to be true, we think, as this evolution has continued into concern with now public sector problems which form a part of what are sometimes referred to as "public management science." The point to stress, we believe, is that all three sectors -- military, private sector and public management science -- were strengthened as further methodological developments occured with the movement into this new "public management science" sector.

Up to now, the emphasis has been on a use of our tools and concepts "in the service of management," so to speak.[1] Although such a course has been a fruitful one, we do not think that it can be considered as controlling in all situations and, in any case, we do not think it represents "the end of the line."

[1] In fact, under the leadership of Gene Woolsey, this has become a guiding principle in the editorial policies of Interfaces. See, e.g., R. E. D. Woolsey [22].

In fact, we shall argue that a next phase in the growth of management science can, and should, be directed to the evaluation and control of management. The trick will be to accomplish this in a way that continues in the same spirit as before -- viz., effecting these extensions in ways that will strengthen rather than weaken what has already been achieved by management science "in the service of management" (military, private enterprise, and public service varieties).

Some of the steps that have already been taken will constitute the main focus of this presentation. How this further evolution in our discipline can also be used to establish contact with approaches of the social sciences that have also been concerned with such evaluations -- e.g., the cost/benefit studies of economists and/or the large-scale social experiments designed by sociogists and others for evaluating programs in areas like education, etc., -- will also be of interest. During this discussion, however, we should also continue to bear in mind the differences in emphasis that come from our concern with management and managed entities.

As indicative of some of these differences, we may here, for the moment, distinguish between "pure predictions" and "control predictions." The former may be secured, for example, from commonly employed social science applications of statistical methods to ascertain various kinds of social regularities or laws. The latter, i.e., the control predictions, are concerned with uses of statistics (and like analytical devices) to ascertain how such regularities may be altered.[1]

---

[1] Charles Christenson has called our attention to Karl Popper's distinction between historical phophecy and social engineering in his discussion of the social sciences in [18], where similar distinctions are made.

in a similar vein we may also distinguish between "research for discovery" and "research for invention" wherein the latter is more closely associated with the "control prediction" approach that we have just described. Note, for instance, that the usual social science distinction between descriptive (including positive) and normative analysis is not really apt here, since research for invention is susceptible to tests for validity, generalizability, etc., -- and in an objective manner -- just as is the case in "research for discovery."[1]

[1]Without attempting to push the matter to a point of possible conflict, it does seem fair to say that it will usually be prudent to inquire what might be done to change the phenomena being studied, as is always necessary in "research for invention," even when "research for discovery" and/or "pure prediction" are primary research objectives. See, e.g., Charles Perrow [17].

There are, however, differences that should be noted. For instance, in research for invention we may want to focus our analyses on only subsets of the available data instead of simply examining all of the data for their possible value as in discovery research. The results from the analyses of these subsets for invention research may then perhaps be applied to other data, as we shall shortly illustrate, in an attempt to see how operations of an entire group of entities may be altered or improved. As a case in point we might, for instance, study subsets of data for entities that have improved their energy consumption efficiencies. This might be done to enable us to ascertain how to improve the operations of other entities or else enable us to estimate wastes and inefficiencies associated with not doing this. Such an approach would require recourse to models and methods that will differ in important respects from those that are currently being employed in econometric studies in U.S. energy problems and the same is und[illegible] of social science modeling in other areas as well.[2]

[2]Further discussion may be found in Charnes, Cooper and Schinnar [9].

2. SELECTION OF TOPICS AND RELATIONS TO COMPREHENSIVE AUDIT

We have now outlined the course we propose to follow, including the contacts that might be made with other sciences -- along with some of the kinds of methodological and conceptual alterations that this might require. We will also do well to indicate aspects of managerial practice that may be served by the kinds of developments we are considering, and for this we turn to "audit practices" and processes--because of the "3rd party accountability functions"[1] that these processes are supposed to serve.

[1] A discussion of these 3rd party accountability functions and how these relate to accounting and auditing may be found in [13.2] and [13.3].

The term audit is sometimes viewed as synonymous with the CPA function of attesting to financial statements and related representations by management. This CPA kind of audit has the servicing of "3rd parties" as a major function and it therefore carries a corollary responsibility for objectivity and professional canons of care and validation that are of interest to us. Such "financial audits" form only a part of the kinds of audits that are now being practiced, however, and among the class of such wider possibilities we shall single out those that we have elsewhere referred to as "comprehensive audits."[2]

[2] See [12].

The term "comprehensive" seems appropriate because this type of audit extends the process of objective appraisal to <u>all</u> aspects of management. By this we mean that the auditor assumes responsibility for the aspect of management that is

designated for audit as well as the way the audit is conducted. The auditor does not stop short with the attest function under comprehensive audit. Instead, he also assumes full responsibility for the report that is rendered, as well as the 3rd party groups to be serviced by the audit process. Finally, instead of restricting audit examinations to financial transactions (and their representations in financial reports) such comprehensive audits may extend to examination and appraisal of any of the following aspects of management and organization behavior:

1. Propriety of
   a. Objectives pursued
   b. Methods used
2. Effectiveness in
   a. Stating objectives
   b. Attaining objectives
3. Efficiency of performance as measured by
   a. Benefits received
   b. Resources utilized.

For the sake of concreteness, we should observe that such audits are not merely constructs from our imaginings. They are presently employed by the U.S. General Accounting Office and they may also be found in the internal audit practices of certain large multinational enterprises.[1] Moreover the value of these kinds of audits are now being recognized by other governmental units, both foreign and domestic, and they are likely to continue to grow as part of the internal control and corporate governance reforms that are presently being considered by many private enterprises.[2] In any case, we shall use the possible needs and

---

[1] See Churchill, et al. in [12].

[2] Including the extension of the audit function that is represented by the Peer Review audits of CPA firms themselves that are discussed in [19].

opportunities of such comprehensive audits to guide our discussion since via its 3rd party orientation, this kind of audit also points in the new directions that we want to explore for management science.

In this paper our focus will be on recently developed methods for objectively assessing various kinds of efficiency[1/] that are pertinent to

[1/] We use this term in the sense indicated above (where it is distinguished from effectiveness and propriety).

the activities of not-for-profit entities. Hence we first ought to say a word about the issues of propriety and effectiveness, which form highly important parts of a widened management accountability that is now beginning to be recognized. On the subject of "propriety" we should probably note that there is a tendency on the part of auditors to use this term to cover situations which range from criminal activities of the grossest sort down to trivial discrepancies in records or the way they have been maintained. This tends to underemphasize the importance of the former and to overemphasize the importance of the latter. Somewhere within these two extremes, moreover, there is a need for perspective that is also often missing. After all, some of the improprieties of yesterday have become the proprieties of today and the former may even have been the precondition for progress from one to the other. How and in what form such improprieties may be ingredients of social progress is a subject that is still only poorly understood and hence we need to approach this in-between area with some tolerance and perspective.

Having said the above, however, one must allow for the fact that until "propriety" can be assumed, it is extremely difficult to bring the issues

of effectiveness or efficiency into perspective. Witness, for instance, the fact that in thousands of pages of testimony dealing with the conduct of activities in sensitive Federal government agencies like the CIA and FBI there is scarcely any discussion of effectiveness or efficiency. These, more positive, aspects of agency operations must (or should) be brought into prominence sooner or later, however, and an unmet challenge of our time is how to control and improve the operations of these kinds of agencies.

[1] Further discussion and references may be found in Churchill and Cooper [11].

The difficult part of the question of effectiveness lies, of course, in the choice (and statement) of objectives. This has been a prime concern of research in the area of decision theory including the more recent multi-attribute extensions for use on the multiple objective decision problems that normally confront not-for-profit enterprises.[2/] Something more seems

[2/] See Keeney and Raiffa [14].

to be needed, however, in moving from the sphere of individual choice to public accountability including due allowance for degrees of possible inconsistency and the adaptations that are likely to be required in the continually changing sets of social values where many not-for-profit entities now operate.

Experiments in mixing a variety of approaches would seem to be in order. As part of a consortium concerned with developing a national goals accounting system, for instance, we are utilizing an approach that we refer to as "goal focusing."[3/] In this approach the elements of "goal programming"[4]

[3/] See the discussion of "goal focusing" in [10]

[4/] See Charnes and Cooper [3].

are combined with "efficient point"[1] considerations so that a choice is effected from among the latter set. In this way the goal programming approach focuses on only a subset of the entire set of efficient points. Note, however, that the closest efficient point may be associated with solutions that are further from the goals than would be the case if other (non-efficient) points were designated.

In the development of these "goal focus" approaches we are trying to connect so-called "quality-of-life" measurements with their economic costs and consequences as part of a national goals accounting system. Hence the needs and approaches utilized by economists need to be kept in mind. The characteristic approach in formal economics is one which proceeds by maximizing a "utility function" under a "budget constraint." An advantage of this approach is that it restricts attention to a relatively small subset of the entire set of efficient-point possibilities while, at the same time, allowing a study of tradeoff possibilities by reference to pricing alternatives. The goal focus approach also restricts attention to a subset of the efficient points but does this in a way that relates the resulting solution to the goal sets while also yielding the tradeoffs in terms of dual evaluators associated with deviations from these goals.

The latter, i.e., the goal sets and their relations to each other, are obtained in various ways which include, e.g., the responses that one might get from interview or questionnaire data such as might be secured from household panels. Many of the goals that are pertinent to quality-of-life dimensions relate to off-market considerations. Hence it seems unwise to retain market price related behavior as the center of all of the analyses. The goal focus approach relaxes this requirement, but in a way that maintains contact with the efficient point considerations of economic analysis so that, inter alia, sensitivity and trade-off analyses can

---

[1] See the discussion in Chapter IX of [4].

be conducted in terms of the "real" economic and other costs which are pertinent for public policy applications. The point is that these costs and tradeoffs are more fundamental than any utility theoretic formulation and all of them need to be accommodated in one systematic analysis such as the goal focus approach is designed to provide via (a) a focus on only a subset of the efficient points and (b) summary measures of the tradeoff possibilities via the optimal dual variables associated with deviations from the goal.[1]

[1] These topics and others are treated in detail in a forthcoming consortium research volume entitled The Production of Well Being which will be published by the National Planning Association under the editorship of N. Terleckyj. See [10] See also Schinnar [21].

There are, of course, many questions that remain and these include the kinds of efficiency estimates and adjustments that might be made and how these might be objectively validated. Other aspects of the topic of efficiency will occupy us in the remainder of this paper, which will be developed as follows. First, we shall indicate what we mean by efficiency and how we propose to measure it. Then we shall apply it to data secured from Program Follow Through, a recently completed large-scale social experiment in U.S. public school education. This will provide an opportunity to distinguish further between what we shall refer to as "program efficiency" and "management efficiency", in which we may visualize the latter as representing the extent to which management takes advantage of the opportunities provided by the former. Finally, we shall attempt to provide a summing up and we shall also provide an appendix that constitutes a start toward the kinds of further statistical developments that are needed in dealing with these kinds of issues.

3. A Model for Efficiency Measurement[1]

To initiate our discussion in this section we may observe that it is customary to measure efficiency as a ratio of wanted outputs to valued inputs both in non-negative measure. The value of this ratio is at most unity and hence we may write

$$1 \geq \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \geq 0 \tag{1}$$

$$\text{with} \quad j = 1, \ldots, n$$

$$\text{and} \quad u_r, v_i > 0, \quad \forall\, r, i$$

to represent the kinds of ratios that we will be considering to measure the efficiency of each of $j = 1, \ldots, n$ decision making units (DMUs). Here the DMUs will be associated with sites in various school districts with the y_{rj} and x_{ij} (all positive) representing, respectively, agreed upon outputs and inputs such as we shall shortly be discussing.[2]

Because the observed y_{rj} and x_{ij} are all assumed to be positive and because the weights u_r, v_i are also be restricted to positive choices, the non-negativity imposed on the ratio in (1) will always be satisfied. Hence it will be eliminated from further explicit attention. The problem of selcting the u_r, v_i for rating any DMU in an objective manner is then resolved by introducing an optimizing principle as follows:

---

[1] This section draws heavily on research done in collaboration with E. Rhodes as reported in [7] and [20].

[2] Other entities might also be used, of course, including internal organization units, etc. See pp. 430-431 in [7].

$$\text{maximize} \quad h_o = \frac{\sum_{r=1}^{s} u_r y_{ro}}{\sum_{i=1}^{m} v_i x_{io}}$$

$$\text{subject to}$$

$$(2) \qquad 1 \geq \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}}$$

$$\text{with } j = 1, \ldots, n$$

$$\text{and } u_r, v_i > 0 \quad \forall\ r, i.$$

The resulting values of the u_r, v_i are then determined entirely from the x_{ij}, y_{rj} data in terms of this model. Here the y_{ro} and x_{io} represent the data for one of the j = 1, ..., n DMUs that has been singled out for efficiency evaluation so that evidently the rating is to be effected relative to the other DMUs in the set. That is, the result is a rating of relative efficiency. Finally, because the ratio in the functional also appears in the constraints we have max $h_o = h_o^* \leq 1$ and $h_o^* = 1$ only if DMU_o is efficient.

As formulated in (2) we appear to be required to deal with optimization of a linear fractional functional under linear fractional constraints. Because we want to focus on conceptual clarity, we will not pursue this topic in detail but we may say that this problem can be replaced by an ordinary linear programming equivalent.1/ Moreover, this linear

1/ See [7].

programming problem and its dual provide not only the wanted measures of efficiency but also the marginal productivities (from the efficient DMUs) and these values can be used to construct the efficient production possibility surface. Thus, in addition to obtaining the wanted efficiency measures, a new simultaneous estimation method for obtaining extremal relations from empirical data is also provided.

The immediately preceding remarks will not be pursued further in this paper because we want to center all of our remaining attention on these efficiency measures. They are introduced here only to help us interpret the results of the optimization in (2). This we may do with reference to the illustrative data of Table 1, which is supposed to represent a situation in which 3 DMUs produce the same single unit of output. The data in the rows for x_1 and x_2 represent the amounts of the inputs utilized by DMU_1, DMU_2 and DMU_3, respectively, in the production of this same one unit of output.[1]

[1] We may also think of this situation as the <u>input</u> amounts <u>per unit output</u> of these 3 DMUs.

Evidently, DMU_2 is not as efficient as DMU_1 and hence cannot be characterized as efficient (i.e., it cannot have $h_o^* = 1$) since DMU_1 has produced the same amount of output (one unit) with one unit less of the first input (i.e., we have $x_{11} = 2$ vs. $x_{12} = 3$) and no more of the second input.[2]

[2] **Formal definitions of the efficiency concepts being used here are provided in [5] and [20] where they are also related to other concepts such as Pareto efficiency in economics.**

TABLE 1

An Illustration of DMU (= Managerial) Efficiency

DMU No. / Input	1	2	3
x_1	2	3	4
x_2	2	2	1

In fact, an application of the simplex method to the data of Table 1 -- see [ ] -- produces an optimal value for DMU_2 of

(3.1) $$h_o^* = 6/7$$

along with optimal weights,[1/]

(3.2) $$u^* = 1,\ v_1^* = 1/6 \text{ and } v_2^* = 1/3.$$

[1/] These are the associated optimal dual variables, as described on pp. 438-440, in [ ].

That this result satisfies all constraints for (2) in this case is readily verified by using the data of Table 1 to obtain

(3.3)
$$\frac{1\ u^*}{2v_1^* + 2v_2^*} = 1$$
$$\frac{1\ u^*}{3v_1^* + 2v_2^*} = \frac{6}{7}$$
$$\frac{1\ u^*}{4v_1^* + 1v_2^*} = 1\ .$$

As a byproduct of our calculation, we may observe that both DMU_1 and DMU_3 are characterized as efficient. The $h_o^* = 6/7$ for DMU_2 (which will also appear in the functional) means that under efficient production only 6/7 of the amount of each input utilized by DMU_2 should have been required. Alternatively, the reciprocal value could have been employed to mean that 7/6 units of output could have been secured in place of the one unit actually realized from the input amounts utilized by DMU_2.

Figure 1 portrays the situation geometrically. Here P_2 corresponds to the data for DMU_2 and P_1 and P_3 similarly correspond to the data for DMU_1 and DMU_3 when the axes are interpreted as providing coordinate values for the input variables recorded in Table 1. The contraction associated with $h_o^* = 6/7$ brings P_2 into coincidence with P_2' where the latter lies on the "unit isoquant" — i.e., the isoquant segment represented by

$$(4.1)\qquad \{(x_1,x_2) : \frac{1}{6} x_1 + \frac{1}{3} x_2 = 1;\ 2 \le x_1 \le 4;\ 1 \le x_2 \le 2\} .$$

Turning to the situation depicted for P_4 in Figure 1, a similar analysis would also give $h_o^* = 6/7$ but now the relevant isoquant segment would be

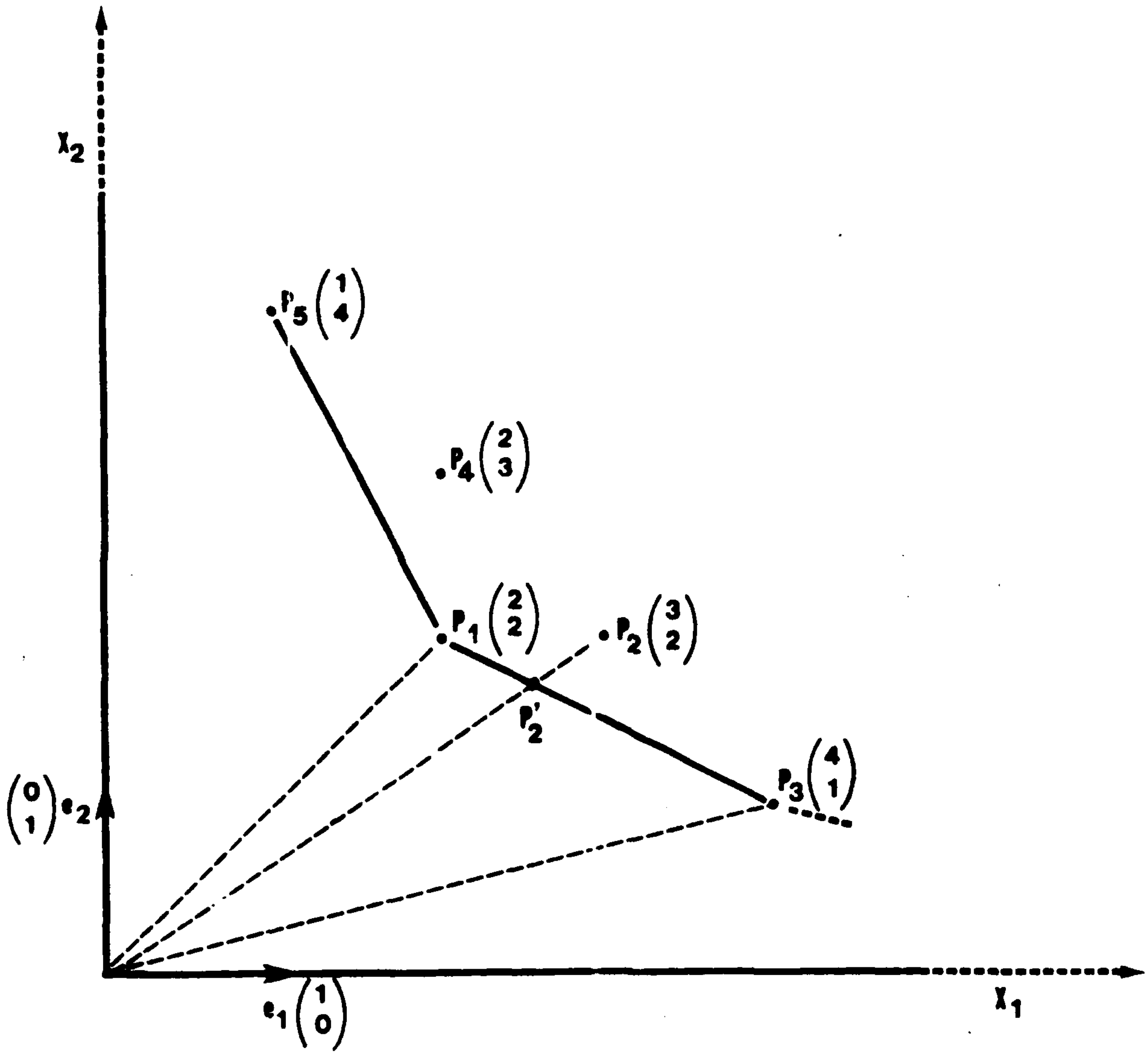
$$(4.2)\qquad \{(x_1,x_2) : \frac{1}{3} x_1 + \frac{1}{6} x_2 = 1;\ 1 \le x_1 \le 2;\ 2 \le x_2 \le 4\} .$$

Thus, although the contraction factor is the same in both cases, the efficient referents differ — since the pertinent isoquant segments are generated from P_5 and P_1 in the case of P_4 and from P_1 and P_3 in the case of P_2.

In a loose sense this difference in the referents provides some advantage for the applications we shall be studying insofar as P_5 and P_1 generate a cone which is more pertinent for P_4 than the cone generated from P_1 and P_3. See [5]. Instead of pursuing the sense of the advantages from this choice of referents, however, we prefer to continue with our discussion of Figure 1 in terms of the results portrayed in (4.1) and (4.2). The values $v_1^* = 1/6$ and $v_2^* = 1/3$ in (4.1) represent the marginal productivities obtained from the efficient isoquant segment connecting P_1 and P_3. These, as already observed, are the efficient productivities obtained from the data for P_1 and P_3. When applied to the inputs utilized by $P_2 = DMU_2$ these productivities give

Figure 1

Efficiency Points and Isoquant



$$(5.1) \qquad 3v_1^* + 2v_2^* = 7/6.$$

Thus we may think of this application of (2) to the data of Table 1 (or Figure 1) as positioning P_2 on the isoquant segment for the output level 7/6 -- which would lie to the northeast and parallel to the segment of the unit isoquant connecting P_1 and P_3 in Figure 1.

In a similar way we could also apply the slope values shown in (4.2) to the data for P_4 to obtain

$$(5.2) \qquad \left(\frac{1}{3}\right) 2 + \left(\frac{1}{6}\right) 3 = \frac{7}{6}$$

so that this same isoquant level would again be attained, but now it would be parallel to the segment between P_1 and P_5.

The interpretation is now staightforward in either case. Our model (2) obtains the productivities from the efficient DMUs and applies them to all other DMUs to estimate what ouputs could have been obtained from the inputs utilized by the latter in each case. In other words, the model (2) applied to the data shown in Figure 1 has given us the ability to construct any part of the entire production function surface which may be pertinent. See [5].

Up to this point we have restricted attention to the single output case so that we could thereby have easy access to the customary production function and related concepts from elementary economics. The numerator of (2) involves multiple outputs, however, and so the concept of a production function must give way to more general concepts for a full interpretation of these ratios.[1/] We want to

[1/] See the discussion of production possibility surfaces in [7]

effect our interpretation in a way that maintains contacts with engineering as well as economics constructs[1] however, and this dictates the way we will now proceed.

It is customary to measure efficiency in engineering when there is a single positive input and a single positive output by the ratio of the output to the input, both measured in the same dimensional units (e.g., in energy units). An energy input is transformed by the process being examined into an output energy in a different form. By the second law of thermodynamics such a transformation cannot increase the amount of energy in the new form over that in the original (input) form.

We generalize this idea to other contexts, or "production processes," by introducing a single "virtual input" and a single "virtual output" obtained by multiplying each input by a weight and summing, and by also multiplying each output by a weight and summing as was done in (2). We do not restrict the dimension of the virtual input and output other than to require that they be the same in order to yield a dimensionless ratio. Further, we require of this dimension for the virtual input and virtual output that it have the property of energy conversion, viz, that the virtual output cannot exceed the virtual input.

To rate the efficiency of a particular DMU -- say DMU_o -- we formulate the extremal principle of choosing the weights so as to satisfy the "energy-degradation" principle for all the DMU's such that the virtual output to virtual input ratio of DMU_o is a maximum. This maximum value represents its efficiency.

All of this is done analytically in (2). It should also be explicitly noted that our efficiency rating, as obtained in (2), is independent of the units in which

[1] Further extension in the economics direction would identify the u_r^* in (2) with efficient "marginal rates of transformation" in a manner parallel to the identification of the v_i^* with the marginal productivity constructs that we have already effected from economics. See Allen [1].

the observed inputs and outputs are measured so long as the units are the same for every DMU. To see that this is so we need only note that if the x_{ij} for a particular DMU are replaced by $x'_{ij} = \rho x_{ij} > 0$ for $j = 1, \ldots, n$ then replacing v_i^* (an optimal value of v_i) by $v'_i = v_i^*/\rho$ and leaving all other u's and v's unchanged we satisfy all contraints and obtain $h'_o = h_o^*$. Thus we must have max $h'_o = h'^*_o \geq h^*_o$. On the other hand, if we could have $h'^*_o > h^*_o$, then, by the inverse transformation, we could obtain an h_o from u'^*, v'^* which would satisfy the original constraints and we would also have $h_o = h'^*_o > h^*_o$. But this contradicts the assumed maximality of h^*_o, and so we must have $h'^*_o = h^*_o$.

Thus we obtain the same efficiency value regardless of the units in which the i^{th} input is measured. A similar argument holds for any other input or output. Our efficiency measure is therefore independent of the units employed in any of the inputs or outputs and hence these units may be chosen for computational (or other) convenience in any fashion.

The background we have now supplied provides access to the well-known computational power and computer code availablities of ordinary linear programming. This means that large numbers of inputs, outputs and DMUs can be handled without any great difficulty. The need for this kind of power and convenience is becoming increasingly apparent in energy (and other) studies which can no longer be adequately addressed by only econometric-statistical techniques and models.[1/]

The interpretations we have supplied made direct contact with both engineering and economics definitions of efficiency.[2/] One can, of course, continue to use these

---

[1/] See, e.g., Manne, et al. [16].

[2/] The u^*_r, v^*_i values have economic significance in their own right as efficient marginal productivities and transformation rates. To emphasize their possible use separate from the above models and developments we refer to them as components of a Productive Efficiency Vector represented by

$$PEV = (u^*_1, u^*_2, \ldots, u^*_s;\ v^*_1, v^*_2, \ldots, v^*_m).$$

in manners that have been customary, although even then new possibilities for application are thereby opened. For instance, one might proceed by other means to estimate the efficient productivities as a basis for determining savings in areas like energy consumption by altering DMU practices for this and other outputs of their operations that might be of interest instead of proceeding, as at present, merely to estimate projected patterns of energy consumption from past practices and efficiency distributions, as is the practice in presently conducted econometric studies.[1] Finally, our approach opens still further possibilities for new and important distinctions such as we shall next undertake to illustrate by distinguishing between "program efficiency" and "management efficiency" for their potential importance for various kinds of evaluations and accountability relations.

---

[1] See, for instance, the Kennedy and other models discussed in [16]. See also [9].

4. DEA: DATA ENVELOPMENT ANALYSIS[1/]

Figure 2, below, will help us to introduce our next topic. We may think of this as a portrayal in which inputs in amounts x have resulted in observed amounts y of a single output (all else held constant) for a collection of DMU's. There is a further distinction in that these DMU's operate under 2 different technologies so that the efficient frontiers for one set are at B while the efficient frontiers of the other set are at A. Evidently even efficient DMU's in B cannot achieve the levels of those in A and even some of the less efficient DMU's operating under A exceed what is attainable under B.

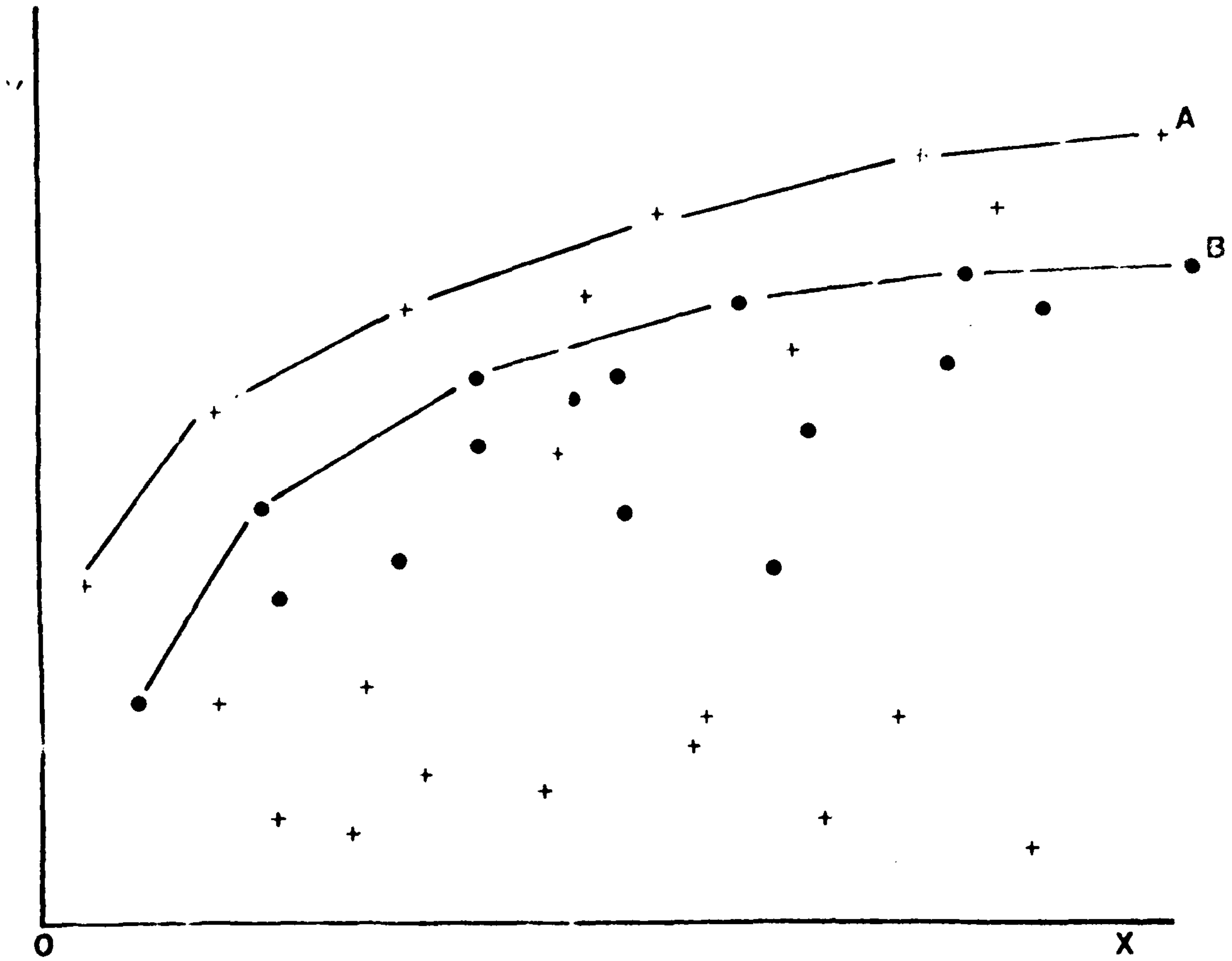
A standard approach via statistical regressions might have concealed some of the alternatives for choice that are present because the observations generally contain a mix of technological possibilities with their utilization. This is the situation we would now like to confront by extending our previous analysis in order to bring it to bear in ways that might help us distinguish between what we shall refer to as "managerial efficiency" and "program efficiency". The latter may be thought of in terms of the kinds of frontiers depicted in Figure 2 while the former may contain inefficiencies resulting from managerial decisions that fail to utilize these opportunities to the full.

"Data Envelopment Analysis" is the name that we shall use for the approach we are suggesting because (a) it first tries to locate the boundaries that envelop the observations as in A and B of Figure 2 and then (b) it brings the observations all up to the envelope that is pertinent in each case. Finally, it imputes any remaining efficiency difference to the respective programs so that in the situation of Figure 2, say, the program

---

1/ This part of our paper draws heavily on work done in collaboration with E. Rhodes as reported in [6]. See also [20].

Figure 2



<u>Legend:</u>

Pluses (+) - Observations for DMU's in A

Dots (●) - Observations for DMU's in B

associated with A would be characterized as more efficient than the program associated with B. Moreover, the numerical value of these efficiencies--measured as the difference in the statistical distributions resulting from the indicated adjustments--is intended to represent the amount of output gain that is attainable by moving the DMU's from B to A.

The above remarks need further clarification which we may obtain by our earlier contrast between "control predictions" and "pure predictions." The latter term, we may recall, is intended to refer to the kind of predictions that one might make from the customary statistical approaches. Fitting regression equations to A and B separately, without otherwise distinguishing between the observations, for instance, might lead to results in which the program associated with B was evaluated as being the better of the two alternatives.

This pure prediction approach could be justified if it were assumed that all DMU's would be permitted to continue to operate at the same levels of managerial efficiency. If, however, new controls to alter this past pattern of behavior could be considered, then an alternate approach might be used in the form of a "control prediction" to justify the choice of A over B.

DEA, Data Envelopment Analysis, is intended for use in the latter situations since, without such controls, the resulting "efficient behavior" predictions will generally be invalid. We may, in fact, think of DEA as providing guidance for program audits which will help to locate sources of inefficiency in order to direct such audit examinations to particular DMU's brought into view by a DEA analysis.[1] A separate decision may then be made concerning whether the projected

---

[1] Further discussion is provided in [2].

efficiencies can be realized and perhaps thereby justify a choice of A rather than B.

To help us distinguish program from managerial efficiency in the different reference sets of DMU's we shall be studying, we now introduce the following extension of (2):

$$\max h_o^\alpha = \frac{\sum_{r=1}^{s_\alpha} u_r^\alpha y_{ro}^\alpha}{\sum_{r=1}^{m_\alpha} v_i^\alpha x_{io}^\alpha}$$

(6) subject to

$$1 \geq \frac{\sum_{r=1}^{s_\alpha} u_r^\alpha y_{rj}^\alpha}{\sum_{i=1}^{m_\alpha} v_i^\alpha x_{ij}^\alpha} \quad ; \quad j=1, \ldots, m_\alpha$$

$$u_r^\alpha,\ v_i^\alpha > 0;\ r=1, \ldots, s_\alpha\ ;\ i=1, \ldots, m_\alpha\ ,$$

where $\alpha=1, 2, \ldots, k$, respectively, indexes the sets which are of interest.

Within each such set we will, of course, have the same efficiency measurement situation as before -- viz., $0 \leq h_o^{*\alpha} \leq 1$ with $h_o^{*\alpha} = 1$ if and only if the DMU being evaluated relative to the α^{th} set of DMUs is efficient. Within each such set we shall assume that we are securing a measure of "managerial efficiency." Only when allowance has been made for the presence of this source of inefficiency will we be in a position to assess the "program efficiency" that is also of interest.

In order to illustrate what is involved we turn to data from "Program Follow-Through", a large-scale experiment in U. S. public school education designed to help disadvantaged children. This program

was intended to further the objectives of a pre-school program known as Project Head Start, by carrying on specially designed programs for disadvantaged children from kindergraten through grade 3. This was done at the 70 selected "school sites" in various parts of tle U.S.A. which are listed in Table A-1. (See first 2 columns of Table A-11 for the numbering that will hereafter be identified with these sites.

An assessment of Program Follow-Through was sought in what was intended as a designed experiment [1/] by the study directors. This was done by selecting a matched pair of Program Follow-Through and Non-Follow-Through sites from which observations were to be secured. These are identified as PFT and NFT, respectively, in columns 1 and 2 of Table A-1. Although we need to observe that the experiment was not carried out successfully in all respects, it nevertheless produced a wealth of data from which we effect a selection as follows.

First, we restrict ourselves only to educational aspects of the program while ignoring other parts of its activities such as supplying medical and dental services, nutritional programs, etc. We also omit any explicit consideration of costs on the supposition that a cost/benefit analysis only becomes pertinent if statistically significant effects are obtained from program elements such as we are considering.

[1/] Or at least as close to such an experiment as one is likely to achieve in contemporary U.S. public school situations

Table A-1

Site Level Distribution of DFA Study Sample

PFT Site #	NFT Site #	Model and Site Name	Region[1]	City Size[2]	PFT[3] Student Pop.	NFT[4] Student Pop.
		Responsive Education Model				
1	50	Berkeley,CA	W	Medium City	99	71
[illegible]	51	Buffalo,NY	NE	Large City	77	27
3	52	Duluth,MN	NC	Medium City	77	79
4	53	Fresno,CA	W	Medium City	48	54
5	54	Lebanon,NH	NE	Rural Area	14	97
6	55	Salt Lake,UT	W	Medium City	36	51
7	56	Tacoma,WA	W	Medium City	51	42
		TEEM Model				
8		Baltimore,MD	S	Large City	99	-
9		Lakewood,NJ	NE	Small City	80	-
10	57	Lincoln,NB	NC	Medium City	96	55
11	58	Wichita KS	NC	Large City	84	36
		Bank Street Model				
12	59	New York,NY	NE	Large City	72	245*
13	60	Philadelphia,PA	NE	Large City	80	37
14		Brattleboro,VT	NE	Small City	20	-
15	61	Fall River,MA	NE	Medium City	39	18
16		Wilmington,DE	S	Medium City	109	-
		DIM Model				
17		New York,NY	NE	Large City	31	-
18	62	E St. Louis,IL	NC	Large City	56	21
19		Grand Rapids,MI	NC	Medium City	103	-
20	63	Racine,WI	NC	Medium City	62	27
21	64	Flint,MI	NC	Medium City	77	66
		BA Model				
22		New York,NY	NE	Large City	43	-
23	65	Philadelphia,PA	NE	Large City	108	27
24		Portageville,MO	NC	Rural Area	47	-
25		Kansas City,MO	NC	Large City	61	-
26		Louisville,KY	S	Large City	90	-
27		Meridian,IL	NC	Rural Area	68	-
		Cognitive Curriculum Model				
28		New York,NY	NE	Large City	52	-
29		Chicago,IL	NC	Large City	18	-
30		Okaloosa Co.,FL	S	Small City	48	-
		Parent Education Model				
31		Philadelphia,PA	NE	Large City	46	-
32	66	Jacksonville,FL	S	Large City	15	53
33	67	Richmond,VA	S	Large City	111	69
34	68	Houston,TX	S	Large City	95	78
		EDC Model				
35		Philadelphia,PA	NE	Large City	112	-
36		Paterson,NJ	NE	Medium City	42	-
		Self-Sponsored Model				
37		Detroit,MI	NC	Large City	43	-
38	69	New York,NY	NE	Large City	20	13
39		Philadelphia,PA	NE	Large City	86	-
40		Portland,OR	W	Large City	45	-
41		San Diego,CA	W	Large City	71	-
		ILM Model				
42		New York,NY	NE	Large City	53	-
		SEDL Model				
43	70	Philadelphia,PA	NE	Large City	86	36
44		Tulare,CA	W	Small City	173	-
		Home-School Partnership Model				
45		New York,NY	NE	Large City	26	-
		California Process Model				
46		Los Angeles,CA	W	Large City	98	-
47		Ravenswood,CA	W	Small City	74	-
48		Lamont,CA	W	Rural Area	27	-
49		San Jose,CA	W	Large City	42	-
				Total Student Pop.	3310	1202

* Pooled Citywide NFT Population Figure for New York City

(Footnotes continued on next page)

Table A-1
(continued)

[1]NE = North Eastern United States

S = Southern United States

NC = North Central United States

W = Western United States

[2]Large City = 200,000 or more

Medium City = 50,000 to 199,999

Small City = 10,000 to 49,999

Rural Area = Less than 10,000

[3]All Data Envelopment Analysis study information refers to the Cohort II-K student population. II-K indicates that this group of students began their Program Follow Through experience in kindergarten. (This was also the only one of three Cohorts which had completed all of the grades from kindergarten through third grade at the time of our study for which site level information was available.) However, due to incomplete statistics along some DEA variable dimensions, some of the Cohort II-K PFT sites were not included in the DEA study. Specifically, Bank Street Model: Rochester,NJ site; EDC Model: Chicago,IL site; and SEDL Model: St. Martin Parish,LA site were excluded from the DEA study student population. The actual Cohort II-K PFT population was 3,367 of which, as noted above, a set of 3,210 students were used in the DEA study. This exclusion of sites also extended to the NFT groups which were similarly reduced to 1,202 students.

[4]Two sets of NFT students groups were created in the original Program Follow Through study. One group was a local student set, usually in the same school system as the subject PFT site. The second group, and the one selected for the DEA study, was a "best matched" group, which may or may not have been located in the same school system or even the same geographical region. The NFT group which most nearly matched the PFT students of a given site along a number of demographic and initial performance dimensions was considered the "best match" for the latter. For several PFT sites the <u>same</u> "best matched" NFT group was used. The much smaller NFT student population total of 1,202 as compared to the PFT student total of 3,210 resulted. See also preceding footnote.

We shall focus on only one of several cohorts from the subjects comprehended in the study. In addition, we shall utilize only the terminal (grade 3) results for this cohort to avoid the additional complications needed to deal with dynamic or transient behavior en route to this terminus. From a set of 11 output measures we select only the following 3 as sufficiently indicative for our purpose:

(7)

y_1 : Total Reading Score as measured by the Metropolitan Achievement Test.

y_2 : Total Mathematics Score as measured by the Metropolitan Achievement Test.

y_3 : Coopersmith Self-Esteem Inventory, intended as a measure of self-esteem.

This y_3 measure, we may note, is directed to non-cognitive growth (or affective behavior) in a dimension that is deemed pertinent to the objectives of this program. Together with y_1 and y_2 this y_3 variable provides a good indication of what is involved in assessing such programs. Note, in particular, that no easily available scheme for weighting the relative importance of these outputs is at hand. Nevertheless, some "overall" measure of program efficiency is wanted in order to enable us to evaluate PFT vs. its NFT alternative and this is to be achieved from data such as are exhibited in Tables A-2 and A-3.

Table A-2

Unadjusted PFT Output Observations

Site #	Total Reading Scores, PHS* Y_1	Total Math Scores, PHS* Y_2	Total Coopersmith Scores, PHS* Y_3
1	54.53	58.98	38.16
2	24.69	33.89	26.02
3	36.41	40.62	28.51
4	14.94	17 58	16.19
5	7.81	6.94	5.37
6	12.59	16.85	12.84
7	17.06	16.99	17.82
8	20.29	30.64	33.16
9	26.13	29.80	26.29
10	46.42	51.59	35.20
11	39.80	37.73	30.29
12	37.84	47.85	25.35
13	26.48	31.36	26.54
14	10.31	10.86	7.47
15	14.39	18.30	14.33
16	32.94	36.03	38.19
17	17.25	20.80	12.07
18	27.55	38.19	20.44
19	41.12	43.80	36.54
20	29.43	47.63	23.34
21	37.46	51.02	27.44
22	19.40	25.18	16.52
23	39.88	47.72	38.97
24	25.72	30.81	16.54
25	24.88	25.27	22.43
26	31.62	40.78	31.16
27	31.31	38.32	25.03
28	21.00	21.30	18.30
29	6.51	7.02	6.16
30	11.64	15.26	15.68
31	12.58	15.90	14.42
32	4.59	6.16	4.99
33	43.76	46.64	39.10
34	32.38	38.55	31.05
35	34.64	45.46	39.22
36	11.52	15.14	13.91
37	15.96	19.21	15.30
38	9.91	12.30	7.22
39	30.44	33.53	29.80
40	22.63	25.24	17.15
41	24.41	27.16	25.30
42	23.11	22.67	17.56
43	21.82	31.45	27.54
44	63.92	79.67	63.11
45	9.47	11.92	8.85
46	33.94	39.18	34.61
47	29.42	35.10	28.42
48	7.70	11.02	9.02
49	12.17	16.03	15.82

* PHS = Per Hundred Students

Table A-3

Unadjusted NFT Output Observations

Site #	Total Reading Scores,PHS* Y_1	Total Math Scores,PHS* Y_2	Total Coopersmith Scores,PHS* Y_3
50	[illegible]9.07	4[illegible].71	27.67
51	9.96	14.34	9.33
52	45.37	51.38	31.61
53	18.23	22.05	17.56
54	59.63	64.41	35.89
55	24.20	28.21	18.74
56	13.53	17.09	15.61
57	28.39	27.65	20.79
58	21.67	26.22	13.66
59	120.17	144.67	88.59
60	15.15	18.04	13.58
61	6.92	7.10	6.35
62	9.35	9.85	7.70
63	13.03	13.40	10.29
64	18.63	24.48	23.13
65	12.28	13.01	9.89
66	16.81	19.72	18.70
67	26.36	28.22	24.46
68	22.85	26.21	28.14
69	8.17	8.70	5.12
70	13.69	14.19	12.99

*PHS = Per Hundred Students

Given these output variables we next turn to a choice of the inputs to be used in our study. From data on the 25 input variables utilized in the Follow-Through study we have selected the following 5 for purposes of our illustration:

(8)

x_1 : Education level of mother as measured in terms of percentage of high school graduates among female parents.

x_2 : Highest occupation of a family member according to a pre-arranged rating scale.

x_3 : Parental visit index representing the number of visits to the school site or with Follow-Through personnel.

x_4 : Parent counseling index calculated from data on time spent with child on school related topics such as reading together, etc.

x_5 : Number of teachers at a given PFT or NFT site.

The data for these variables from both PFT and NFT sites, respectively, are arranged in Tables A-4 and A-5.

Table A-4

Unadjusted PFT Input Observations

Site #	Education Level of Mother, PHS*	Occupation Index, PHS*	Parental Visit Index, PHS*	Counseling Index, PHS	Number of Teachers
	X_1	X_2	X_3	X_4	X_5
1	36.13	16.24	48.21	49.69	9
2	29.26	10.24	41.96	40.65	5
3	[illegible]3.12	11.31	38.19	35.03	9
4	24.96	6.14	24.81	25.15	7
5	11.62	2.21	6.85	6.37	4
6	[illegible]1.88	4.97	18.73	18.04	4
7	32.64	6.88	28.10	25.45	7
8	20.79	12.97	54.85	52.07	8
9	34.40	11.04	38.16	42.40	8
10	61.74	14.50	49.09	42.92	9
11	52.92	11.67	39.48	39.64	5
12	36.00	10.15	37.80	39.52	5
13	39.20	10.80	41.04	41.12	7
14	14.6	2.88	9.64	11.14	3
15	4.29	5.42	21.45	17.27	5
16	27.25	14.17	56.46	55.26	9
17.	22.63	4.43	15.40	15.00	2
18	28.00	7.61	28.73	27.04	9
19	[illegible]3.56	13.70	53.04	49.85	7
20	[illegible]5.42	9.05	29.69	31.74	4
21	[illegible]1.57	10.08	39.34	40.57	6
22	16.34	5.84	20.89	22.10	4
23	44.28	14.14	56.70	52.27	11
24	[illegible]9.74	6.43	24.20	25.66	3
25	24.40	8.05	33.42	31.29	7
26	41.40	11.70	44.01	46.35	7
27	27.20	9.38	37.80	31.55	4
28	23.92	7.12	25.58	29.01	3
29	10.62	2.55	10.10	9.09	4
30	12.48	6.14	23.13	22.46	6
31	19.32	5.89	24.01	24.74	6
32	6.30	1.93	7.11	7.68	4
33	46.62	14.65	65.71	57.49	10
34	38 95	12.82	47.02	48.92	9
35	[illegible]	15.56	53.98	50.29	6
36	31.08	6.26	22.18	21.96	4
37	19.35	6.68	22.61	23.31	4
38	11.20	3.08	9.90	10.06	2
39	34.40	11.61	41.79	41.79	5
40	35.55	6.48	21.69	21,69	6
41	30.53	9.30	35.50	35.14	8
42	25.44	7.10	26.81	26.23	3
43	26.66	11.43	41.36	44.63	6
44	39.79	22.49	84.77	76.12	11
45	8.32	3.64	12.92	13.13	2
46	59.78	13.52	48.80	49.69	15
47	39.22	10.06	37.00	38.33	4
48	3.28	3.18	13.12	12.71	5
49	7.14	5.29	23.10	19.06	8

*PHS = Per Hundred Students

Table A-5

Unadjusted NFT Input Observations

Site #	Education Level of Mother, PHS*	Occupation Index, PHS*	Parental Visit Index, PHS*	Counseling Index, PHS*	Number of Teachers
	X_1	X_2	X_3	X_4	X_5
50	68.16	12.28	33.58	34.64	15
51	11.88	3.59	13.41	13.82	8
52	55.30	11.53	36.73	35.78	6
53	16.20	7.02	26.94	26.30	9
54	82.45	15.52	45.00	44.23	13
55	15.81	6.93	23.91	23.61	7
56	4.65	5.50	20.91	23.39	5
57	41.25	8.41	26.23	25.24	10
58	10.44	5.22	17.10	18.93	3
59	139.65	35.03	119.56	130.83	22
60	16.28	4.81	18.20	18.98	5
61	12.06	2.59	8.74	8.17	5
62	4.20	2.64	9.89	11.25	2
63	19.44	3.83	12.87	13.23	5
64	28.38	8.91	30.95	33.33	8
65	13.50	3.61	15.60	12.39	4
66	23.32	7.10	24.96	28.56	22
67	27.60	9.38	32.29	34.01	20
68	11.70	10.53	37.67	43.60	8
69	4.68	1.85	6.22	5.46	5
70	10.44	4.82	17.13	18.21	9

*PHS = Per Hundred Students

Given the preceding data we try next to disentangle any managerial inefficiencies that might otherwise appear as program inefficiencies. This is done by assigning $\alpha = 1$ to PFT and $\alpha = 2$ to NFT in (6) and separately calculating the respective $h_o^{*\alpha}$ values that are exhibited in Table A-6. In each case the efficient members of the set generate an "envelope," or efficiency frontier, which we shall refer to as α-envelopes. *Thus, e.g.,* $\alpha = 1$ *refers to the* envelope for PFT and $\alpha = 2$ refers to the envelope for NFT which we shall also label as the "1-envelope" and "2-envelope," respectively.

One question we may now ask is whether the managerial efficiencies are the same for both sets of observations or whether the results of a PFT-NFT comparison might be clouded by differences in this very important dimension. Ideally, this should have been allowed for in the statistical design but, of course, it would be difficult to obtain the needed data on efficiencies in advance. However, we can at least make allowances and/or effect adjustments for this along lines like the following.

For instance, we might calculate the proportion of managers who achieve values of $h_o^{*\alpha} = 1$ under $\alpha = 1$ and $\alpha = 2$, respectively. This gives

$$(9) \qquad P(\alpha = 1) = \frac{m_e(1)}{m(1)} = \frac{17}{49} = \frac{51}{147}$$

where $m_e(1)$ is the number of PFT DMU's with h_o^{*1} values of 1.0 and $m(1)$ is the number of DMU's in PFT. Similarly for NFT we obtain

$$(10) \qquad P(\alpha = 2) = \frac{m_e(2)}{m(2)} = \frac{8}{21} = \frac{56}{147}$$

where $m(2)$ refers to the total number of DMU's in NFT and $m_e(2)$ refers to the number on the efficiency frontier.

Table A-6

PFT and NFT Program Specific α-Envelope Efficiency Values

PFT Site #	h_0^{*1} Efficiency Value	NFT Site #	h_0^{*2} Efficiency Value
1*	1.00	50	0.95
2	0.90	51	0.92
3	0.98	52*	1.00
4	0.90	53	0.87
5*	1.00	54*	1.00
6	0.90	55*	1.00
7	0.89	56*	1.00
8	0.91	57	0.92
9	0.87	58*	1.00
10*	1.00	59	0.92
11	0.98	60	0.98
12	0.97	61	0.88
13	0.86	62*	1.00
14	0.98	63	0.96
15*	1.00	64	0.91
16	0.95	65	0.97
17*	1.00	66	0.92
18*	1.00	68*	1.00
19	0.95	69*	1.00
20*	1.00	70	0.94
21*	1.00		
22*	1.00		
23	0.96		
24*	1.00		
25	0.97		
26	0.93		
27*	1.00		
28	0.94		
29	0.84		
30	0.90		
31	0.83		
32	0.90		
33	0.94		
34	0.85		
35*	1.00		
36	0.80		
37	0.94		
38	0.94		
39	0.91		
40*	1.00		
41	0.94		
42	0.94		
43	0.87		
44*	1.00		
45	0.89		
46	0.90		
47*	1.00		
48*	1.00		
49*	1.00		

* Denotes a site with an efficiency value of "1"

Proceeding now in a somewhat informal manner, we can see that the managers in the two programs, PFT and NFT, have about an equal likelihood of being on their respective program referenced efficiency frontiers. The differences between $P(\alpha=1)$ and $P(\alpha=2)$ are not statistically significant so that an implication of the above results is that both PFT and NFT site managers or "decision-makers" are drawn from the same managerial efficiency pool or population. [illegible]atively speaking there appear to be just as many managers in PFT as in NFT who, within the limits of their program constraints, operate their sites efficiently.

The above probability measures provide insight only into the relative location of DMU's in PFT and NFT with respect to their boundaries. Questions involving differences between the distribution of efficiency values in the two groups are not addressed by the above calculation. Thus to check and perhaps extend our understanding of these results, while still staying with relatively simple measures, we consider a comparison between the mean efficiency value differences in the two sets. Again using the information contained in Table 2, we compute a PFT mean efficiency as $\bar{h}_o^{*1} = 0.946$ and an NFT mean efficiency value as $\bar{h}_o^{*2} = 0.958$. Thus, NFT has a slightly higher mean efficiency but again the difference is not statistically significant.

Having arrived at these results we are at the point of choosing whether to elect a "pure prediction" or a "control prediction" course along the lines that we indicated when discussing Figure 2. Proceeding along the latter lines we might arrive at a situation such as the one shown in Figure 3. Here the envelopes for PFT and NFT cross over from one part of the diagram to another which means that PFT is more efficient in the region to the right and NFT is more efficient in the left-hand portion of the diagram.

Analyses could be centered on each of the latter segments, of course, but something more might be wanted. We might, in particular, want an overall assessment to enable us to choose between the two programs and for this we might proceed as follows. First, we might replace the originally observed output values y^1_{rj} and y^2_{rj} with new values $\hat{y}^1_{rj}$ and $\hat{y}^2_{rj}$ which result when the originally observed values are adjusted in a manner that brings them onto their respective envelopes.[1] Similarly, we can replace the original x^1_{ij} and x^2_{ij} input values in Tables A-4 and A-5 by new values $\hat{x}^1_{ij}$ and $\hat{x}^2_{ij}$ that are derived from their efficiency frontiers. Then we can derive a new envelope which we shall refer to as the "inter-envelope," via

$$\max.\ \hat{h}_o = \frac{\sum_{r=1}^{s} u_r \hat{y}_{ro}}{\sum_{i=1}^{m} v_i \hat{x}_{io}}$$

subject to

$$(11) \qquad 1 \geq \frac{\sum_{r=1}^{s} u_r \hat{y}^1_{rj}}{\sum_{i=1}^{m} v_i \hat{x}^1_{ij}} \; ; \quad j=1, \ldots, m_1$$

and

$$1 \geq \frac{\sum_{r=1}^{s} u_r \hat{y}^2_{rj}}{\sum_{i=1}^{m} v_i \hat{x}^2_{i}} \; ; \quad j=1, \ldots, m_2$$

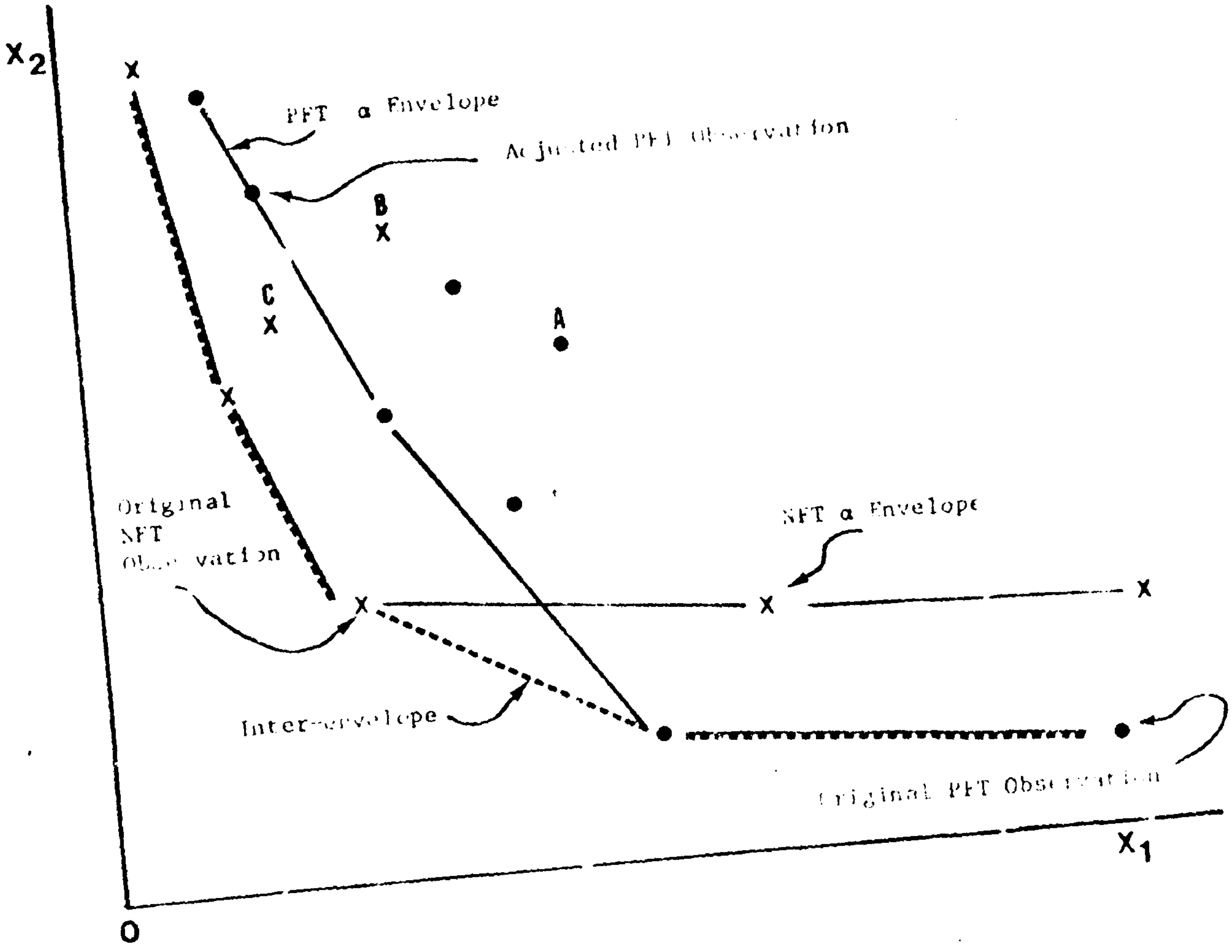
[illegible] as before,[2] all variable values are constrained to be positive.

---

[1] The manner of doing this is given in [7] and [10].

[2] Supra (2)ff.

Figure 3



Legend:

● = PFT Observations

X = NFT Observations

——●—● = α Envelope for PFT

——X—X = α Envelope for NFT

.... = Inter-envelope

Application of (11) produces the results presented in Table A-7, as well as the remaining results needed for the explication of Figure 3. Note that we are now assumint that all managers are operating efficiently, by which we mean that they will always move to the boundaries indicated by their respective envelopes. We similarly assume that they will move to any new boundaries that may become available when their present confinement to either PFT or NFT boundaries only is relaxed, and this is what the inter-envelope comparison is intended to provide.[1] Notice, for example that the inter-envelope is always as efficient as either the $\alpha = 1$ or the $\alpha = 2$ envelope, and, in fact, it coincides with them when one or the other is more efficient.

All that remains now in effecting our evaluation is to select a suitable statistic that will reflect how the $\hat{h}_0^*$ values are distributed for $\alpha = 1$ and $\alpha = 2$ relative to their potential values when the present boundaries are removed. This is provided by the Kullback-Leibler statistic[2] for which we get

$$I\ (\alpha = 1) = 2.40226 \times 10^{-4}$$

and

$$I\ (\alpha = 2) = 0.03684 \times 10^{-4}.$$

This means that the distribution for the adjusted PFT observations, taken all together, are more distant from the inter-envelope than is the case for the distribution of the NFT observations.[3] In other words PFT fails the overall efficiency test at least for the cohort and the variables that we have examined.

---

[1] More precisely, we are assuming that these managers can be brought onto the boundaries -- while, of course, we are also providing estimates of the losses (or wastes) that will be experienced when this is not done. This is to say that our adjustments are enforcing these assumptions for all DMU's although, of course, we need not always proceed in this manner. We may in fact restrict our adjustments only to subsets of the DMU's and/or make only partial adjustment of designated inputs or outputs when desired.

[2] See [14].

[3] "On the Statistical Distribution of DMU Efficiency Measures" by A. Charnes, W. W. Cooper and E. Rhodes in [8] provides a canonical development which is illustrated by a direct PFT/NFT comparison. This comparison, which does not utilize the inter-envelope, yields the same conclusion -- viz., the data and criteria used in this paper do not show PFT to be more efficient than NFT in any statistically significant sense.

Table A-7

Inter-Envelope Efficiency Values

PFT Site #	h_0^* Efficiency Value	NFT Site #	h_0^* Efficiency Value
1	0.92	50*	1.00
2*	1.00	51*	1.00
3	0.94	52*	1.00
4*	1.00	53*	1.00
5	0.93	54*	1.00
6*	1.00	55	0.99
7	0.99	56*	1.00
8*	1.00	57*	1.00
9	0.98	58*	1.00
10	0.92	59*	1.00
11*	1.00	60	1.00
12*	1.00	61*	1.00
13	0.99	62*	1.00
14	0.95	63*	1.00
15*	1.00	64*	1.00
16*	1.00	65*	1.00
17*	1.00	66*	1.00
18*	1.00	67*	1.00
19	0.99	68	0.99
20*	1.00	69*	1.00
21*	1.00	70*	1.00
22*	1.00		
23	0.99		
24*	1.00		
25*	1.00		
26	0.99		
27*	1.00		
28*	1.00		
29	0.99		
30*	1.00		
31	0.99		
32*	1.00		
33	0.99		
34	0.98		
35*	1.00		
36*	1.00		
37	0.94		
38	0.99		
39*	1.00		
40	0.95		
41	0.99		
42*	1.00		
43	0.99		
44*	1.00		
45	0.99		
46*	1.00		
47*	1.00		
48*	1.00		
49*	1.00		

*Denotes a site with an efficiency value of "1"

Of course, this need not end the matter since, as we have already noted, a facet-by facet comparison[1] might yield situations in which PFT is superior to

[1]The basic sets assoicated with the simplex (and dual method) calculations as described in [4] and [5] will make it easy to distinguish the various facets.

NFT. In addition, the further possibilities associated with the broken line segment shown for the inter-envelope in Figure 4 might also be noted. It will be observed that this segment is not coincident with either of the original envelopes. The reason is that this segment contains "mixture" possibilities that are not apparent from either envelope alone. That is, it represents situations in which the PFT and NFT programs might be combined to yield still further possibilities for improvement. In other words, our analysis has not only yielded the wanted program comparisons but it has also yielded a flexible way of effecting more detailed (facet by facet) comparisons as well as entirely new program possibilities in terms of one analytical approach.

5. CONCLUSION

Appropriate conclusions will be drawn from the discussion of this paper at the Hawaii TIMS meetings and entered in a revised version of this paper. The latter will be sent to interested persons in response to requests sent to either author. In the meantime we also herewith supply an appendix which provides a characterization which is canonical for the kinds of statistical distribution involved along with an illustration of their use when direct PFT vs NFT comparisons instead of the comparisons effected through the inter-envelope of Figure 3 are wanted.

ADDENDUM

We might better point the concluding section of this paper into possible new directions as follows. After the above plenary session paper was presented at the TIMS Meetings in Honolulu, Hawaii, Alan Hoffman[1] of IBM observed to us

[1] See [20] for comments on Dr. Hoffman's contributions to the efficiency studies of M. J. Farrell at an earlier date.

that the concepts in our paper might be used to study comparative social-economic systems, in which event one might find that the Russian (socialist) system was better but failed to show this because it was more poorly managed than the U.S. (capitalist) system. Our reply was that if such a study were conducted, its authors had better be aware of the difference between control predictions and pure predictions made elsewhere in this paper.

An opportunity to explore these ideas further was provided in a post-meetings Radical Workshop on Organization Design organized by Arie Lewin and Kenneth MacKenzie on the island of Maui. One whole morning being set aside at this Workshop for the organization design implications of our efficiency measures suggested a more modest route to us than the undertaking mentioned by Dr. Hoffman. In particular, we were aware that a great deal of the small group work undertaken by Dr. MacKenzie[2] in the field of organization design has often been concerned with multiple

[2] See, e.g., K. D. MacKenzie, A Theory of Group Structure -- Vol. I, Basic Theory and Vol. II, Empirical Tests (New York, Gordon & Breach, 1976).

outputs. To be sure, such analyses are generally confined to only a very few output measures - such as (1) time to complete a task with different kinds of organization arrangements and (2) degree of satisfaction with the work by persons in various parts of each type or organization -- and almost never involve more than a few input measures such as time and number of persons. On the other hand, these data do provide a start and, of course, our efficiency measure should provide a basis for extension into more complex input and output considerations in the future.

A systematic basis for further progress in this work may thus be provided that could then be extended to the study of efficiency-effectiveness interactions -- and perhaps other new dimensions for research might thereby also be opened. The point on which we can now conclude is that program vs. managerial efficiency or even organization vs. managerial efficiency under different programs are not the only possibilities. What is required for the immediate future is some meaningfully measurable way of identifying different sources of efficiency, and laboratory experimentation along the dimensions of small group/organization design research certainly offers an inviting prospect for such possibilities.

BIBLIOGRAPHY

[1] Allen, R. G. D., Mathematical Analysis for Economists (New York: The MacMillan Co., 1939).

[2] Charnes, A. and W. W. Cooper, "Auditing and Accounting for Managerial Efficiency and Program Efficiency in Not for Pofit Entities," Plenary Session Paper, American Accounting Association Meetings in Honolulu, August 24, 1979.

[3] ______ ______, "Goal Programming and Multiple Objective Optimizations, Part I," European Journal of Operational Research 1, No. 1, January 1977, pp. 39-54.

[4] ______ ______ Management Models and Industrial Applications of Linear Programming (New York: John Wiley & Sons, Inc., 1961).

[5] ______ ______ Managerial Economics -- Past, Present and Future," Journal of Enterprise Mangement I, No. 1, 1978, p. 15.

[6] ______ ______ ______, "A Data Envelopment Analysis Approach to Evaluation of the Program Follow Through Experiment in U.S. Public School Education," Management Science (submitted).

[7] ______ ______ and E. Rhodes, "Measuring the Efficiency of Decision Making Units," European Journal of Operational Research 2, No. 6, Nov. 1978, pp. 429-444.

[8] ______ ______ ______, "On the Statistical Distributions of DMU Efficiency Measures," Research Memo. (Austin, Texas, Center for Cybernetic Studies, 1978).

[9] ______ ______ and A. Schinnar, "Transforms and Approximations in in Cost and Production Function Relations," The Bell Journal of Economics (submitted).

[10] ______ ______ ______ and N. Terleckyj, "A Goal Focusing Approach to Analysis of Trade-Offs Among Household Production Outputs," in N. Terleckyj, ed., Production of Well-Being (Washington: National Planning Association, forthcoming).

[11] Churchill, N. C. and W. W. Cooper, "Comprehensive Auditing." Paper presented to the Institue of Chartered Accountants of Ontario, Canada, June 12, 1978.

[12] ______ ______, J. G. San Miguel, J. Pond and V. Govindarajan, "Comprehensive Audits and Some Suggestions for Research," Symposium on Auditing Research, II (Dept. of Accountancy, University of Illinois at Urbana-Champaign, 1977).

[13] Cooper, W. W. and Y. Ijiri, eds., Eric Louis Kohler -- Accounting's Man of Principles (Reston, VA, Reston Publishing Co., 1979).

[13.1] Brasfield, K. A., "Development of GAO Auditing Standards and Their Relevance to the Practice of Public Accounting," pp. 157-167.

[13.2] Churchill, N. C. and W. W. Cooper, "Auditing and Accounting -- Past, Present and Future," pp. 225-239.

[13.3] Cooper, W. W. and Y. Ijiri, "Accounting and Accountability Relations," pp. 191-207.

[14] Keeney, R. L. and H. Raiffa, Decisions with Multiple Objectives: Preferences and Value Tradeoffs (New York: John Wiley & Sons, Inc., 1976).

[15] Kullback, S. Information Theory and Statistics (New York: John Wiley & Sons, 1959).

[16] Manne, A. S., R. G. Richels and J. P. Weyant, "Energy Policy Modeling: A Survey," Operations Research 27, No. 1, Jan-Feb., 1979, pp. 1-36.

[17] Perrow, Charles "Three Types of Effectiveness Studies" in P. S. Goodman, J. M. Pennings and Associates, New Perspectives on Organizational Effectiveness (San Francisco: Jossey-Bass, Publishers, 1979).

[18] Popper, Karl R., "Historical Prophecy versus Social Engineering" and "Technological Approach to Sociology," in The Poverty of Historicism (New York: Harper & Row Inc. 1961).

[19] Public Oversight Board, Annual Report 1978-79 (New York 10020: Public Oversight Board, American Institute of Certified Public Accountants, 1978).

[20] Rhodes, E., Data Envelopment Analysis and Related Approaches for Measuring the Efficiency of Decision Making Units with an Application to Program Follow Through in U.S. Public School Education (Pittsburgh: Carnegie-Mellon University, School of Urban and Public Affairs, 1978). Ph.D. Thesis.

[21] Schinnar, A. P., Models for Eco-Demographic Planning and Forecasting (Pittsburgh: Carnegie-Mellon University, School of Urban and Public Affairs, 1976), Ch. 6, Ph.D. Thesis.

[22] Woolsey, R. E. D. "A Triennial Editorial Policy," Interfaces 6, No. 4, August 1976.

APPENDIX

ON THE DISTRIBUTION OF DMU EFFICIENCY MEASURES

by

A. Charnes*
W. W. Cooper**
E. Rhodes***

*The University of Texas
**Graduate School of Business, Harvard University and
School of Urban and Public Affairs, Carnegie-Mellon University
***State University of New York at Buffalo

This research was partly supported by NSF Grant No. SOC76-15876 "Collaborative Research on the Analytical Capabilities of a Goals Accounting System," and by Project NR 047-021, ONR Contract N00014-75-C-0616 with the Center for Cybernetic Studies, The University of Texas, and ONR Contract N00014-76-C-0932 at Carnegie-Mellon University School of Urban and Public Affairs. Reproduction in whole or in part is permitted for any purpose of the U.S. Government.

Note: This appendix is drawn from a Research Memorandum of the same title released in February of 1978.

Introduction

The discussion in the text was directed to distinctions between program efficiencies and managerial efficiencies and ways in which the two might be separated for purposes of program evaluation and control. These topics are sufficiently novel and complex so that it seemed best to focus on them without diverting attention to still further possibilities.

One such additional alternative that needs to be considered involves the case in which adjustments to achieve improved efficiency in management practices cannot (or will not) be made. Then the evaluation must be made along the lines that we have characterized as the "pure prediction" situation (in which it is assumed that management practices will continue with the same mix of efficiency levels as before).

Undoubtedly still other alternatives might need to be considered. Witness, for instance, the situation for energy studies where, as described in the text, only adjustments for energy consumption and related uses might be of interest on the assumption that other parts of the relevant DMU operations will subsequently be adjusted to these more efficient methods of energy utilization. In any case the unusual character of the statistical distributions being considered makes it desirable to establish canonical references. If, for example, we wish to employ the Kullback-Leibler (or "minimum discrimination information") measures of discrepancy, as in the text, we must compare on the basis of distributions that are mutually absolutely continuous.[1] A simple way to affect this is to estimate for each example the parameters of a "canonical distribution" from a class which has this mutual property.

---

[1]See [14], pp. 6 ff.

The remainder of this appendix is therefore pointed toward the development of such a canonical class. We do this in a way that relies on fundamental properties of the gamma distribution since the latter appears appropriate to the non-negative measurement scales of the inputs and outputs.

Background

The diagrams shown in the Figures at the end of this appendix, which are drawn from the data of Table A-6 in the text, will serve as background for the developments we shall now introduce. We may think of these as figures derived in the following manner.

Each program sample of DMU's determines a vector of sample efficiencies, one for each DMU. These efficiency ratings are obtained from a ratio of linear functionals which is optimized over a polyhedral constraint set that is the same for every DMU.

To make the sense of this more concrete we write our sample data for a particular program in the form

$$(X^1, Y^1), \ldots, (X^j, Y^j), \ldots, (X^n, Y^n)$$

where each X,Y pair corresponds to a vector of input and output observations for one of j=1, ..., n DMU's in that program. The efficiency for the jth DMU is then obtained via

$$h_j^* = \max \frac{u^T Y^j}{v^T X^j}$$

$$\text{subject to} \quad \begin{aligned} u^T Y^1 - v^T X^1 &\le 0 \\ &\cdots\cdots \\ u^T Y^j - v^T X^j &\le 0 \\ &\cdots\cdots \\ u^T Y^n - v^T X^n &\le 0 \end{aligned}$$

with u^T, $v^T > 0$ wherein the T superscripts indicate the transpose of the corresponding column vector.

Now let s be the slack vector with non-negative component $s_1, \ldots, s_j, \ldots, s_n$, which can be associated with the above inequalities to bring them into equation form. Then we can observe (i) that there are a finite number of extreme points (u^k, v^k, s^k)-- $k=1, \ldots, K$ --which span this constraint set and (ii) the optimal h_j (X,Y) is attained for some (u^k, v^k, s^k). We therefore have the following theorems at our disposal:

<u>Theorem</u>: h_j^* (X,Y) = 1 if $s_j^k = 0$ for some extreme point (u^k, v^k, s^k).

<u>Theorem</u>: In every sample h_j^* (X Y) = 1 is attained for at least one j.

Thus the <u>population</u> *density function* for efficiency, arising from the <u>population</u> from which the program sample of n DMU's and the sample (X^j, Y^j) for each j is drawn , must have a part at efficiency = 1. We shall refer to this as the "atomic part" and assume in the development which follows that the efficiency density function is a mixture of a continuous density and the atomic density at t=1. See Figures at the end of this appendix.

<u>Density Characterizations and Developments</u>

We proceed now to choose a class of <u>continuous</u> density functions from the following considerations: First, the vectors X^j, Y^j are assumed to be vectors of sample averages, one average for each input and output of each DMU. These averages, <u>a priori</u>, are <u>non-negative</u>. If we assume that each average is from a gamma distribution of the form $g_{p\lambda}(t) = \frac{1}{\Gamma(p)} \lambda^p t^{p-1} e^{-\lambda t}$ then so is each $u_r^k y_{rj}$ and each $v_i^k x_{ij}$ where these u_r^k and v_i^k are components of the positive vectors u^k and v^k. Assuming further that these averages are

statistically independent of one another[1] and that they have the same ultimate "decay" rate λ, then a numerator N (say $\sum_r u_r^k y_{rj}$) and its denominator D (say $\sum_r v_i^k p_{ij}$) specifying an efficiency rating are each independently gamma distributed with parameters, (p,λ) and (q,λ), respectively.

Now[2] the density function for N/D is

$$h(t) = \begin{cases} \frac{\Gamma(p+q)}{\Gamma(p)\,\Gamma(q)} \quad t^{p-1}(1+t)^{-p-q} & , t > 0 \\ 0 & , \text{elsewhere} \end{cases}$$

where

$$\Gamma(s) \equiv \int_0^\infty t^{s-1}e^{-t}dt \ , \qquad , \Gamma(s) = (s-1)! \text{ for } s \text{ integer.}$$

But, further, our constraints require that $N/D = \sum_r u_r^k y_{rj} / \sum_i v_i^k x_{ij} \leq 1$. Thus, the density function $\bar{h}(t)$ for N/D conditional on $N/D \leq 1$ is

$$\bar{h}(t) = \begin{cases} h(t)/\int_0^1 h(s)ds, & 0 \leq t \leq 1 \\ 0 \qquad , & \text{elsewhere} \end{cases}$$

We note that the mode of the numerator is at $\hat{x}_N = (p-1)/\lambda$, and that of the denominator is at $\hat{x}_D = (q-1)/\lambda$, so that the mode of $\bar{h}(x)$ -- and also that of $h(x)$ -- is at

$$\hat{x} = (p-1)/(q+1),$$

i.e., for large p and q is approximately at the ratio of the modes. From the empirical fact that most operations are better than say 75% or 80% efficient, we can expect from the functional form for $\bar{h}(x)$ that (i) the values of p and q will be quite large, (ii) that $p-1 < q+1$, and (iii) that p and q are about the same magnitude.

---

[1]This does not imply that their parameters are not functionally related.
[2]See M. Dwass [12], p. 287 and also p. 243.

Because of these large powers, p and q, it is not convenient to carry out integrations to determine $\tilde{h}(x)$. Thus to estimate p and q from sample data, we have fit the sample relative frequency function $h(x)$ (non-atomic part) as follows:

We take the sample mode to be $\hat{x}$. Substituting $p-1 = \hat{x}(q+1)$, we have

$$\tilde{h}(x) = x^{(q+1)\hat{x}}(1+x)^{-(q+1)(1+\hat{x})} \quad , \quad 0 \le x < 1 .$$

Then, considering two values x^1, x^2, we get

$$\frac{\tilde{h}(x^1)}{\tilde{h}(x^2)} = \frac{h(x^1)}{h(x^2)} = \left(\frac{x^1}{x^2}\right)^{(q+1)\hat{x}} \left(\frac{1+x^1}{1+x^2}\right)^{-(q+1)(1+\hat{x})}$$

Therefore,

$$\ln\left[\frac{\tilde{h}(x^1)}{\tilde{h}(x^2)}\right] = \ln\left[\frac{h(x^1)}{h(x^2)}\right]$$

$$= (q+1)\left\{\hat{x}\ \ln\left(\frac{x^1}{x^2}\right) - (1+\hat{x})\ \ln\left(\frac{1+x^1}{1+x^2}\right)\right\} .$$

In our case, since (a) we have more data near the mode, and (b) we need the values near x=1 to fit well, we have chosen $x^1 = \hat{x}$ and x^2 = the midpoint of the next larger histogram interval.

For example, from the PFT data histogram at the end of this appendix, we chose $\hat{x}=0.94$, $x^1=x'=x=0.94$, $x^2=0.98$. The results were: $q+1 \cong 342$, $p-1 \cong 321.48$. As mentioned before, it is critical to have precise evaluations of the logarithms in the curly brackets and on the left. Here the former had the value 0.0012 and we therefore obtained $q+1=(0.4101)/0.0012$.

Evidently our chosen form $\bar{h}(x)$ can reasonably represent the non-atomic portion of our observed sample histograms for PFT/NFT and Total PFT and NFT Efficiencies. We have thus achieved our end of obtaining a usable, easily specified, parametric class of continuous density functions to represent efficiency populations.

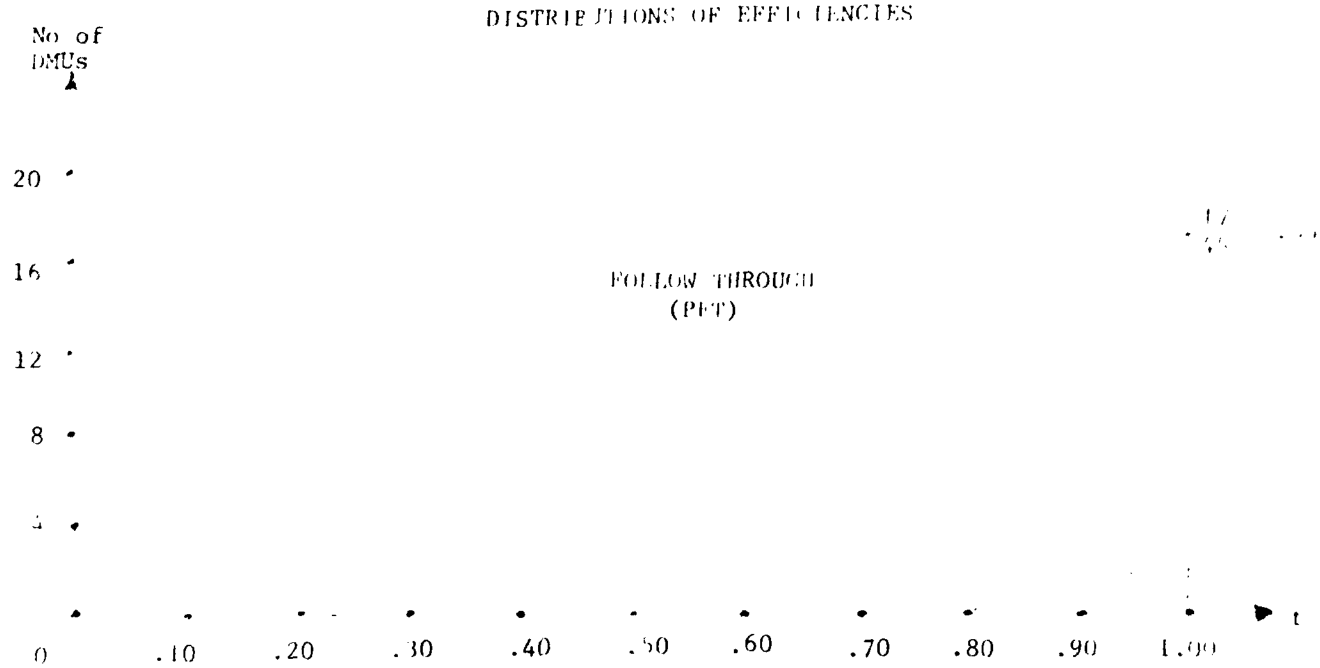
To make comparisons between programs by, say, the Kullback-Leibler "divergence"[1/] of these canonical density functions, we would estimate the s, q parameters for the NFT program just as we have done for the PFT program. We do not complete this numerical calculation here, however. Instead we simply display the formula for determining the Kullback-Leibler divergence in the following form,

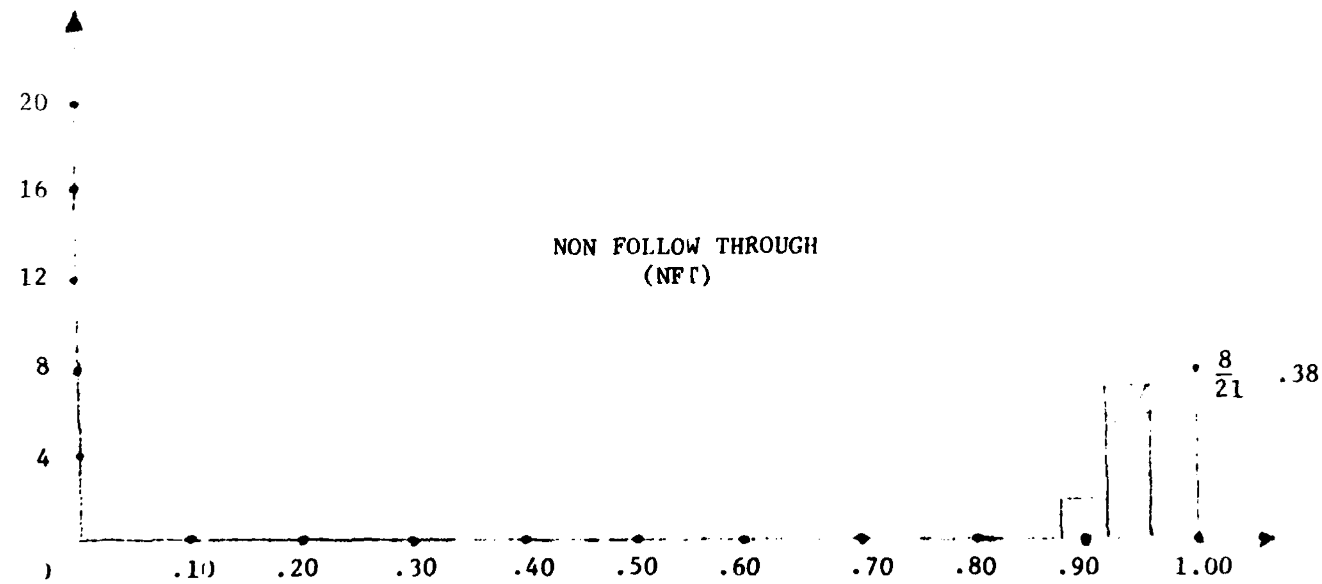
$$J(1:2) = \int_0^1 [(1-f_1)\bar{h}^1(t)-(1-f_2)\bar{h}^2(t)] \ln \left[\frac{(1-f_1)\bar{h}^1(t)}{(1-f_2)\bar{h}^2(t)}\right] dt + (f_1-f_2)\ln \frac{f_1}{f_2}$$

where f_1, f_2 represent the proportions of efficient points for PFT, NFT respectively — e.g., 0.35 and 0.38-- as in the following Figures, and $\bar{h}^1(t)$ and $\bar{h}^2(t)$ are the continuous density parts for PFT and NFT, respectively, as determined from our canonical distribution. From this formula we have evidently guaranteed the mutual absolute continuity properties that this measure requires.

1/ See Kullback [14] pp 6 ff.

DISTRIBUTIONS OF EFFICIENCIES





Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Center for Cybernetic Studies The University of Texas at Austin	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

Management Science Relations for Evaluation and Management Accountability

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

5. AUTHOR(S) (First name, middle initial, last name)

A. Charnes and W W. Cooper

6. REPORT DATE	7a. TOTAL NO OF PAGES	7b. NO OF REFS
May 1979	51	21

8a. CONTRACT OR GRANT NO N00014-75-C-0616 b. PROJECT NO c. NR047-071 d.	9a. ORIGINATOR'S REPORT NUMBER(S) Center for Cybernetic Studies Research Report CCS 341 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Office of Naval Research (Code 434) Washington, DC

13. ABSTRACT

Management accountability as an added dimension for management science research is examined from the standpoint of possible uses in some of the newer "comprehensive auditing" approaches to propriety, effectiveness and efficiency evaluations of management and organization behavior. Attention is centered on non-market activities and not-for-profit organizations. "Goal focusing" is examined, for example, as a relatively recent extension of goal programming for use in effectiveness evaluation and as an alternative to utility theoretic approaches in national goals accounting systems designed to deal with programs or objectives involving numerous kinds of off-market activities. The bulk of the paper, however, is devoted to a new mathematical programming model for deriving analytic representations of extremal frontiers or envelopes from empirical data and for measuring the efficiency of not-for-profit entities. An illustrative application to a recently completed large-scale social experiment in educating disadvantaged children in the U.S. public schools is used to show how distinctions may also be drawn between "program efficiency" and "management efficiency." The appendix develops a canonical form for the types of statistical distributions involved. It also provides a beginning for dealing with statistical properties of the extremal relations obtained by applying these kinds of mathematical programming models to observational data.

Unclassified

Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Accountability						
Audit and evaluation						
Comprehensive audit						
Goal focusing						
Efficiency measurement						
Decision making units (DMU's)						
Data envelopment analysis (DEA)						
Ratio programming						
Statistical distributions						
Kullback-Leibler statistic						

DD FORM 1 NOV 65 1473 (BACK)

S/N 0102-014-6800

Unclassified

Security Classification

A-31409